Appendix 1: Outcome definition

In a similar manner to the definition of mortality, the specific definitions of the outcomes: recurrence of stroke, bleeding events, and readmission for heart failure are provided.

Recurrence of stroke

Presence of repeat hospitalisation for ischaemic stroke (as defined below) following index hospitalisation for ischaemic stroke. Admission to an inpatient unit or emergency department with any of the following primary ICD10 codes:
- ❖ I63* Cerebral infarction.

Excluding:

Any admissions with the admission code:
- ❖ 81-Transfer of any admitted PATIENT from other Hospital Provider other than in an emergency (NHS Data Dictionary).

Bleeding events

Any of the following primary ICD10 codes:
- ❖ K25.0 Acute gastric ulcer with haemorrhage;
- ❖ K25.2 Acute gastric ulcer with both haemorrhage and perforation;
- ❖ K25.4 Chronic or unspecified gastric ulcer with hemorrhage;
- ❖ K25.6 Chronic or unspecified gastric ulcer with both hemorrhage and perforation;
- ❖ K26.0 Acute duodenal ulcer with hemorrhage;
- ❖ K26.2 Acute duodenal ulcer with both hemorrhage and perforation;
- ❖ K26.4 Chronic or unspecified duodenal ulcer with hemorrhage;
- ❖ K26.6 Chronic or unspecified duodenal ulcer with both hemorrhage and perforation;
- ❖ K27.0 Acute peptic ulcer, site unspecified, with hemorrhage;
- ❖ K27.2 Acute peptic ulcer, site unspecified, with both hemorrhage and perforation;
- ❖ K27.4 Chronic or unspecified peptic ulcer, site unspecified, with hemorrhage;
- ❖ K27.6 Chronic or unspecified peptic ulcer, site unspecified, with both hemorrhage and perforation;
- ❖ K28.0 Acute gastrojejunal ulcer with hemorrhage;
- ❖ K28.2 Acute gastrojejunal ulcer with both hemorrhage and perforation;
- ❖ K28.4 Chronic or unspecified gastrojejunal ulcer with hemorrhage;
- ❖ K28.6 Chronic or unspecified gastrojejunal ulcer with both hemorrhage and perforation;
- ❖ K92.0 Hematemesis;
- ❖ I60* Subarachnoid hemorrhage;
- ❖ S06.6 Traumatic subarachnoid hemorrhage;
- ❖ I62.0 Nontraumatic subdural hemorrhage;
- ❖ S06.5 Traumatic subdural hemorrhage;
- ❖ I61* Intracerebral hemorrhage;
- ❖ I62.1 Nontraumatic extradural hemorrhage;
- ❖ I62.9 Nontraumatic intracranial hemorrhage, unspecified;
- ❖ K92.1 Melena;
- ❖ I85.0 Oesophageal varices with bleeding;
- ❖ I98.3 Oesophageal varices with bleeding in diseases classified elsewhere.

OR

Administration of any of the following medications:

- ❖ Dried prothrombin complex;
- ❖ Fresh frozen plasma;
- ❖ Idarucizumab;

OR

Administration of any of the following procedures (OPCS4):

- ❖ G20.1 Fibreoptic endoscopic coagulation of bleeding lesion of oesophagus;
- ❖ G46.2 Fibreoptic endoscopic coagulation of bleeding lesion of upper gastrointestinal tract;
- ❖ X33.2 Intravenous blood transfusion of packed cells.

Readmission for heart failure

Inpatient admission or emergency department admissions 2+ days in length which meet the following criteria:

- ❖ A primary ICD10 code for HF defined as:
    - ◆ I50* Heart failure;
    - ◆ I11.0 Hypertensive heart disease with heart failure;
    - ◆ I13.0 Hypertensive heart and renal disease with (congestive) heart failure;
    - ◆ I13.2 Hypertensive heart and renal disease with both (congestive) heart failure and renal failure.

OR

- ❖ A secondary ICD10 code for HF as (defined above) AND IV Furosemide administered within admission or within 1 day following admission

Excluding

- ❖ Any admission that is less than 48 hours during which one of the following procedures is performed:
    - ◆ Cardioverter defibrillator introduced through the vein (OPCS-4: K59*);
    - ◆ Other cardiac defibrillator (OPCS-4: K72*);
    - ◆ Cardiac pacemaker system introduced through vein (OPCS-4: K60*);
    - ◆ Other cardiac pacemaker system (OPCS-4: K61*);
    - ◆ Other cardiac pacemaker system introduced through vein (OPCS-4: K73*);
    - ◆ Cardiac pacemaker system (OPCS-4: K74*).

Appendix 2: Novel feature definitions

The specific novel feature definitions that are discussed in this paper are provided here. Each line is connected by an 'AND' or 'OR' statement, which can be further connected by an indentation of the same logic.

Computed tomography angiography of cerebral vessels

U212: Computed tomography NEC
AND
(
Z342: Aortic arch
OR
Z35: Cerebral artery OR Z361: Carotid artery NEC
)

Speech disturbances not elsewhere specified

All R47 subcodes combined

Rehabilitation

Z501: Other physical therapy
OR
Z505: Speech therapy
OR
Z507: Occupational therapy and vocational rehabilitation, not elsewhere classified

Hemiplegia

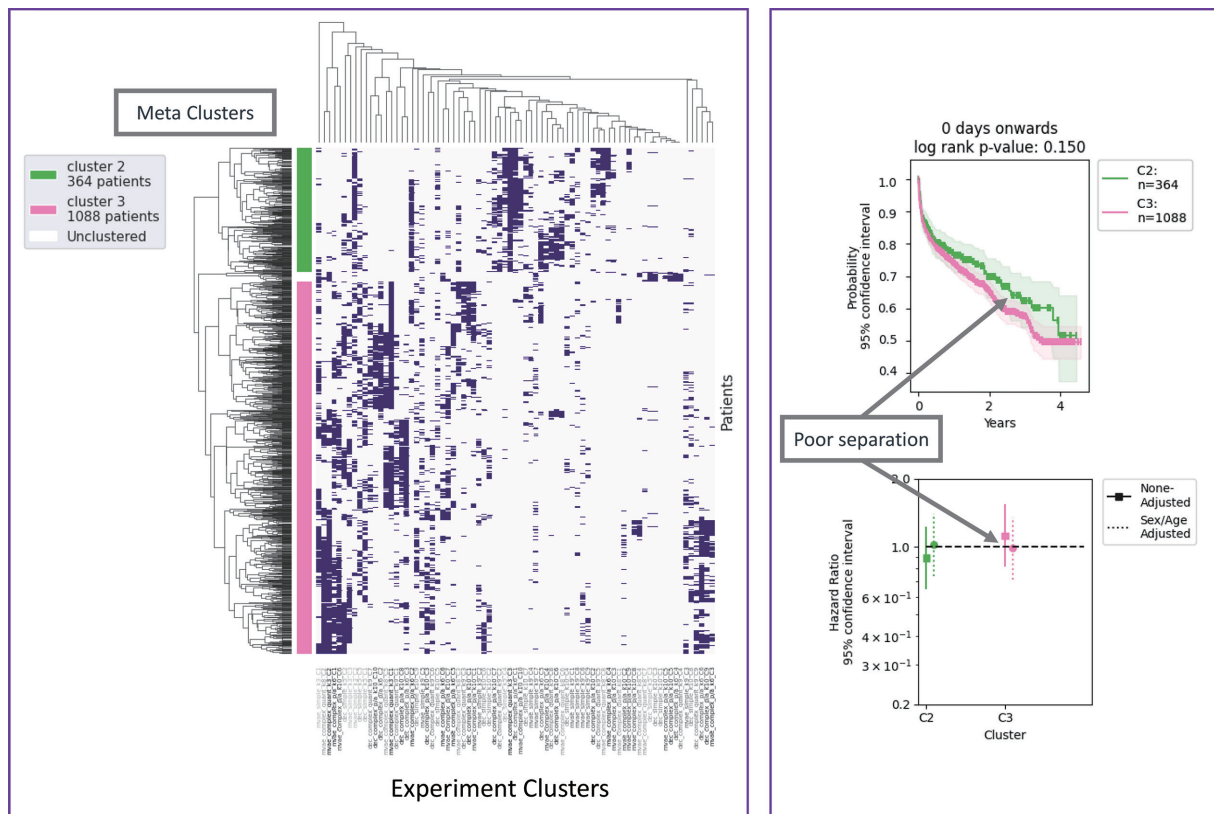All G81 subcodes combined.

Magnetic resonance angiography of cerebral vessels

U211: Magnetic resonance imaging NEC
AND
(
Z342: Aortic arch
OR
Z35: Cerebral artery OR Z361: Carotid artery NEC
)

Appendix 3: Example of poor performance of meta clustering (*Figure S1*)

Experiment C, Oxford University Hospital (OUH) Stroke with k=5 is an example of a meta clustering analysis that did not perform well. In contrast to the example in Figure 6 where we see good separation with distinct blocks of patients in the dendrogram, there is no clear consensus of which clustering the patients should belong. Note that the Kaplan Meier and Cox plots are provided purely for illustrative purposes, meta clustering does not optimise the separation of outcomes.

Figure S1 Experiment C, an example of poor performance of meta clustering (left) with a dendrogram heatmap and (right) survival analysis. Here the patient groups do not show significantly different outcomes.